

Local Multiplicative Schwarz Algorithms for Convection-Diffusion Equations

Xiao-Chuan Cai¹ and Marcus Sarkis²

Department of Computer Science

University of Colorado

Boulder, CO 80309

Abstract

We develop a new class of overlapping Schwarz type algorithms for solving scalar convection-diffusion equations discretized by finite element or finite difference methods. The preconditioners consist of two components, namely, the usual two-level additive Schwarz preconditioner and the sum of some quadratic terms constructed by using products of ordered neighboring subdomain preconditioners. The ordering of the subdomain preconditioners is determined by considering the direction of the flow. We prove that the algorithms are optimal in the sense that the convergence rates are independent of the mesh size, as well as the number of subdomains. We show by numerical examples that the new algorithms are less sensitive to the direction of the flow than either the classical multiplicative Schwarz algorithms, and converge faster than the additive Schwarz algorithms. Thus, the new algorithms are more suitable for fluid flow applications than the classical additive or multiplicative Schwarz algorithms.

¹ The work was supported in part by the NSF grant ASC-9457534, the National Aeronautics and Space Administration under NASA contract NAS1-19480 while the author was in residence at the Institute for Computer Applications in Science and Engineering, the NSF Grand Challenges Applications Group grant ASC-9217394 and the NASA HPCC Group grant NAG5-2218.

² The work was supported in part by the NSF grant ASC-9406582, and in part by the NSF Grand Challenges Applications Group grant ASC-9217394 and by the NASA HPCC Group grant NAG5-2218.

1. Introduction. In this paper, we present some new overlapping domain decomposition methods for the numerical solution of large, sparse, nonsymmetric and/or indefinite linear systems of equations arising from Galerkin finite element discretizations of elliptic partial differential equations. The new algorithms belong to the family of overlapping Schwarz methods which is a variant of the classical Schwarz alternating algorithm, introduced in 1870 by H. A. Schwarz [22] in an existence proof of elliptic boundary value problems defined in certain irregular regions. This family of methods has attracted much attention in the past few years as convenient and powerful in the solution of partial differential equations, see, especially on parallel machines. For tutorial presentations, see, e.g. [7,23].

We shall focus on linear nonsymmetric and/or indefinite second-order elliptic finite element or finite difference equations. The solution of such problems is an important computational kernel in implicit methods, such as solving systems involving the Jacobian in any Newton-like method used in computational fluid dynamics, [3, 4, 24]. This family of methods is built upon the so-called subdomain mapping operators T_i , which solve the original problem, defined on a domain Ω , approximately in subdomains $\Omega'_i \subset \Omega$ with artificial boundary conditions and zero extensions to $\Omega - \Omega'_i$. The formal definitions of T_i and Ω'_i will be given in the next section. By using these T_i 's as basic building blocks, a family of *polynomial Schwarz algorithms* can be defined. Let N be the number of subdomains and T_0 the coarse space mapping operator. We define

$$T = \text{poly}(T_0, T_1, \dots, T_N)$$

as a multi-dimensional matrix-valued polynomial with variables T_i , and assume that the polynomial satisfies $\text{poly}(0, \dots, 0) = 0$, which simply means that the constant term in the polynomial is zero. It is known that if u^* is the exact solution of the finite element equations then Tu^* can be computed without knowing u^* itself. This is because $T_i u^*$, $i = 0, \dots, N$, can be computed directly from the right-hand side function of the finite element equations. With $g \equiv Tu^*$ as the new right-hand side vector, a new linear system can be introduced as

$$Tu = g$$

and it is not difficult to show that if T is nonsingular then the new linear system gives the desired finite element solution u^* . For each choice of the polynomial poly , a particular Schwarz algorithm is defined. The algorithm is called *optimal* if the condition number, or some other “equivalent measure” for nonsymmetric or indefinite problems, of the operator T is independent of the mesh parameter h and the number of subdomains N . Several such optimal algorithms, such as the additive ($T = \sum_{i=0}^N T_i$) and multiplicative ($T = I - \Pi_{i=0}^N (I - T_i)$) Schwarz algorithms, have been identified. Generally, the additive algorithms have two features among others:

- They converge more slowly than the multiplicative algorithms because of the lack of subdomain-to-subdomain communications within each iteration;
- Their convergence is independent of the ordering and coloring of the subdomains.

The features of the multiplicative algorithms include:

- They are faster in terms of the total iteration number;

- They are not as parallel as the additive algorithms because of the data dependence between overlapping subdomains
- They have a strong dependence on the global ordering and coloring of subdomains especially for convection-diffusion problems.

See the last section of this paper for a detailed discussion on the ordering and coloring issues. To use the multiplicative algorithms efficiently, it is important to color and order the subdomains correctly. However, to obtain the optimal coloring and ordering is difficult in practice especially when the underlying mesh is unstructured and the subdomains are obtained by means of graph partitioning, see, e.g., [3, 10]. For a particular problem and a given subdomain partitioning, it is not impossible to obtain a reasonable subdomain coloring and ordering according to certain practical heuristics, but, in general, especially for unsteady problems where the flow direction changes from time step to time step, it becomes desirable to have algorithms that do not need, or depend less on, manual subdomain ordering and coloring. Extensive discussions on the effects of ordering and coloring of nodes or elements, in the context of iterative and direct sparse matrix computations, can be found in many research papers, see, e.g., [1, 9, 15, 17]. Some of the ideas and techniques can also be applied, with certain modifications, to the coloring and ordering of overlapping subdomains. We will not consider these techniques in this paper, since our interest is in automating the construction of the preconditioner.

In this paper, we shall identify some overlapping Schwarz algorithms, which we call the *local multiplicative Schwarz* methods. The new algorithms are not only optimal but also have convergence rates that are:

- better than that of the additive Schwarz method;
- not sensitive to the coloring and global ordering of the subdomains, nor the flow direction;
- more parallel than the multiplicative Schwarz algorithm.

Our basic idea is to use the multiplicative Schwarz algorithm only locally between those pairs of overlapping subdomains for which we have effective techniques to determine the flow direction without any global operations. We use additive techniques to handle the global communication between pairs of subdomains and the coarse level preconditioning.

The paper is organized as follows. In Section 2, we define our model elliptic problem, its discretization and the overlapping partitioning of the finite element mesh. In Section 3, we introduce and analyze the new local multiplicative Schwarz algorithms for symmetric and general nonsymmetric problems. In the last section of the paper we provide some numerical examples regarding the performance of the new algorithms, as well as some comparisons with the classical additive and multiplicative Schwarz algorithms.

2. Model problems and subdomain partitioning. Let Ω be an open, bounded polygonal region in R^d , $d = 2$ or 3 , with boundary $\partial\Omega$. We consider the homogeneous Dirichlet boundary value problem

$$(1) \quad \begin{cases} Lu(x) &= f(x) & \text{in } \Omega, \\ u(x) &= 0 & \text{on } \partial\Omega. \end{cases}$$

Here the elliptic operator L has the form $Lu(x) = -\nabla \cdot (\nabla u) + 2\beta(x) \cdot \nabla u + c(x)u$. All the coefficients are, by assumption, sufficiently smooth and the right-hand side $f \in L^2(\Omega)$. We assume that the equation has a unique solution in $H_0^1(\Omega)$. Let (\cdot, \cdot) denote the usual $L^2(\Omega)$ inner product and $\|\cdot\|$ or $\|\cdot\|_{L^2}$ the corresponding norm. The weak form of equation (1) is: Find $u \in H_0^1(\Omega)$ such that

$$(2) \quad b(u, v) = (f, v), \quad \forall v \in H_0^1(\Omega).$$

The bilinear form $b(u, v)$ is defined by

$$b(u, v) = \int_{\Omega} \nabla u \cdot \nabla v dx + \int_{\Omega} (\beta \cdot \nabla u) v dx + \int_{\Omega} \nabla \cdot (\beta u) v dx + \int_{\Omega} (\tilde{c} u v) dx.$$

Here $\tilde{c}(x) = c(x) - \nabla \cdot \beta$. In addition to the following bilinear form

$$a(u, v) = \int_{\Omega} \nabla u \cdot \nabla v dx,$$

which is used as the usual energy inner product in $H_0^1(\Omega)$ with norm defined by $\|u\|_a = (a(u, u))^{1/2}$, we also use two other bilinear forms

$$s(u, v) = \int_{\Omega} (\beta \cdot \nabla u) v dx + \int_{\Omega} \nabla \cdot (\beta u) v dx$$

and $c(u, v) = (\tilde{c} u, v)$, which correspond to the skew-symmetric and zeroth order parts of L , respectively. It is easy to verify that

$$s(u, v) = -s(v, u), \quad \forall u, v \in H_0^1(\Omega).$$

Following Dryja and Widlund [13], we define a two-level conforming finite element triangulation of Ω . The region Ω is first divided into nonoverlapping subdomains Ω_i , $i = 1, \dots, N$, such that $\bar{\Omega} = \bigcup_{i=1}^N \bar{\Omega}_i$. Then all the subdomains Ω_i , which are assumed to have diameter of order H , are divided into triangular elements of size h . We assume that the union of all of the elements of size h , forms a regular finite element triangulation of Ω . The common assumption, in finite element theory (cf. [8]), that all elements are shape regular is adopted. With such a triangulation, we let $V_h \subset H_0^1(\Omega)$ be the usual piecewise linear continuous finite element space on Ω . To obtain an overlapping decomposition of the domain, we extend each subdomain Ω_i to a larger region Ω'_i , i.e. $\Omega_i \subset \Omega'_i \subset \Omega$. We assume that the overlap is uniformly large and let $V_i \equiv V_h \cap H_0^1(\Omega'_i) \subset V_h$ be the usual finite element subspace defined over Ω'_i , with zero extension to $\Omega - \Omega'_i$. Here uniformly large overlap means that $\text{distance}(\partial\Omega'_i \cap \Omega, \partial\Omega_i \cap \Omega) \geq cH$, where $c > 0$ is a constant independent of H . It is clear that $\bar{\Omega} = \bigcup_i \bar{\Omega}'_i$ and $V_h = V_0 + V_1 + \dots + V_N$.

The finite element discretization of (2) reads as follows: Find $u^* \in V_h$ such that

$$(3) \quad b(u^*, v) = (f, v), \quad \forall v \in V_h.$$

Another key ingredient in the design of optimal domain decomposition preconditioners is the use of at least one global coarse space, which in a way connects the local subdomains

just introduced. A number of coarse spaces have been introduced in the literature, see, e.g., [11, 12]. We shall focus only on a simple one. Let $\Omega_H = \{\tau_i\}$ be a quasi-uniform triangulation of Ω and τ_i one of the triangles with a diameter on the order of H . Ω_H is the coarse grid. Let V_0 be the piecewise linear continuous finite element space on Ω_H . In the analysis part of this paper we assume, for simplicity, that $V_0 \subset V_h$, and that the diameter of the coarse elements τ_i is of the same order as the diameter of the subdomains Ω_i . The theory can easily be extended to the case of a non-nested coarse [2], and to cases with small overlap [14].

In the numerical experiments section, we shall present some cases where the sizes of the subdomains and the coarse elements are of different order.

For each $i = 0, 1, \dots, N$, we define a mapping operator $T_i : V_h \rightarrow V_i$ by

$$(4) \quad b(T_i u, v) = b(u, v), \quad \forall u \in V_h, \quad \forall v \in V_i.$$

These T_i will serve as the basic building blocks of the algorithms to be discussed in the next sections. We shall mention that these T_i 's can also be defined inexactly if we replace the left-hand side bilinear form in (4) by a different bilinear form, which, in some sense is equivalent to $b(\cdot, \cdot)$. Details on inexact Schwarz algorithms can be found in, for example, [5, 7, 23].

3. New algorithms and analysis. In this section, we define the local multiplicative Schwarz algorithms by using the basic Schwarz building blocks T_i defined in the previous section. For each pair of neighboring subdomains, with indices i and j , we define a multiplicative Schwarz operator

$$P_{ij} = I - (I - T_j)(I - T_i).$$

Note that for any $u \in V_h$, $P_{ij}u \in V_i + V_j$, and generally $P_{ij} \neq P_{ji}$, unless Ω'_i and Ω'_j have no common points. Let

$$(5) \quad P = T_0 + \sum P_{ij},$$

where the summation is taken over all possible P_{ij} 's. Let $g_{ij} = P_{ij}u^*$ and $g_0 = T_0u^*$; as mentioned earlier, both can be computed without the knowledge of u^* . With $g \equiv g_0 + \sum g_{ij}$, it can be seen that if the operator P is nonsingular, then the linear system

$$(6) \quad Pu = g$$

has the same solution as that of (3). We shall prove in the remainder of the paper that P is indeed nonsingular and uniformly well-conditioned, and that therefore (6) can be solved by using certain Krylov space type iterative acceleration methods, such as CG or GMRES [21]. We remark that if the bilinear form $b(\cdot, \cdot)$ is symmetric, then the operator P is also symmetric with respect to $b(\cdot, \cdot)$. In other words, the local multiplicative Schwarz operator P is symmetric if both P_{ij} and P_{ji} are included in its definition. Later, in this section, we shall intentionally destroy the symmetry by dropping one of the two terms when solving nonsymmetric problems. Keeping only the terms in the upwind direction makes the algorithm very useful for convection-diffusion equations. Like other upwinding type discretization schemes, we shall also introduce a parameter μ that controls the amount of the upwinding, or artificial diffusion, in the Schwarz preconditioning polynomial.

3.1. Analysis for the symmetric positive definite case. Since the symmetric positive definite case is rather simple, we consider it here. Throughout this subsection we assume that $b(\cdot, \cdot) \equiv a(\cdot, \cdot)$. The full abstract theory of Dryja and Widlund, [13], on the optimal convergence of the additive Schwarz methods cannot be used directly because our subproblem operators P_{ij} are not defined as projections. We summarize the results of the symmetric case in the following theorem.

THEOREM 1. *There exist positive constants c and C , independent of the mesh parameters h and H , such that*

$$c\|u\|_a^2 \leq a(Pu, u) \leq C\|u\|_a^2,$$

for any $u \in V_h$

One of the key facts that we shall use in the proof of the theorem is given in the following lemma due to Dryja and Widlund [13].

LEMMA 3.1 (DRYJA AND WIDLUND[13]). *There exist positive constants c and C , independent of the mesh parameters, such that*

$$c\|u\|_a^2 \leq \sum_{i=0}^N \|T_i u\|_a^2 \leq C\|u\|_a^2,$$

for any $u \in V_h$.

We remark again that since both P_{ij} and P_{ji} are included in the definition of P , the operator P is symmetric with respect to $b(\cdot, \cdot)$. The upper bound of the operator P can be obtained easily, since

$$\|P_{ij}u\|_a^2 \leq C(\|T_i u\|_a^2 + \|T_j u\|_a^2)$$

and

$$\|Pu\|_a^2 \leq C \sum \|P_{ij}u\|_a^2.$$

By using Lemma 3.1, we obtain that $\|Pu\|_a \leq C\|u\|_a$, for any $u \in V_h$ and where $C > 0$ is a constant independent of the mesh parameters. To obtain the lower bound, we note that

$$(7) \quad a(P_{ij}u, u) = \|T_i u\|_a^2 + \|T_j u\|_a^2 - a(T_i u, T_j u).$$

Using the fact that $a(T_i u, T_j u) \leq \|T_i u\|_a \|T_j u\|_a \leq 1/2(\|T_i u\|_a^2 + \|T_j u\|_a^2)$, we have

$$a(P_{ij}u, u) \geq \frac{1}{2} (\|T_i u\|_a^2 + \|T_j u\|_a^2).$$

This gives the lower bound when combined with Lemma 3.1.

We remark that since the operator P is symmetric and positive definite with respect to the inner product $a(\cdot, \cdot)$, the conjugate gradient method can be used. It is obvious that the degree of parallelism of the new method is higher than that of the symmetrized multiplicative Schwarz algorithms. Here we have considered only Poisson's equation; the extension of the algorithm and theory to general variable coefficient symmetric positive definite cases is straightforward.

The analysis for the symmetric case is included above for theoretical interest. In practice, however, we do not believe that this type of upwinding preconditioning would offer much improvement over the classical additive Schwarz method for symmetric positive definite problems. Some numerical examples are included in the last section of the paper.

3.2. Analysis for the general nonsymmetric case. We consider the general nonsymmetric case in this subsection. The techniques are mainly borrowed from Cai and Widlund [5, 6]. Let us begin by summarizing the main results, namely that the operator P is uniformly bounded and its symmetric part, with respect to the inner product $a(\cdot, \cdot)$, is uniformly positive definite, in the following theorem. This theorem provides the optimal convergence of several Krylov space iterative methods, including GCR [16] and GMRES [21] among others.

THEOREM 2. *There exist positive constants H_0 , $c(H_0)$ and C , independent of the mesh parameters h and H , such that if $H \leq H_0$, the operator P is uniformly bounded, i.e.,*

$$\|Pu\|_a \leq C\|u\|_a, \quad \forall u \in V_h,$$

and its symmetric part is uniformly positive definite, i.e.,

$$a(Pu, u) \geq c\|u\|_a^2, \quad \forall u \in V_h.$$

To prove the above theorem, we need a result from Cai and Widlund [6] regarding the optimality of the additive Schwarz preconditioner.

LEMMA 3.2 (CAI AND WIDLUND[6]). *There exist positive constants H_0 , $c(H_0)$ and C , independent of the mesh parameters, such that if $H \leq H_0$*

$$\left\| \sum_{i=0}^N T_i u \right\|_a \leq C\|u\|_a, \quad \forall u \in V_h,$$

and

$$\sum_{i=0}^N \|T_i u\|_a^2 \geq c\|u\|_a^2, \quad \forall u \in V_h.$$

We next present a number of useful lemmas before giving the proof of the main theorem later in this subsection. The following lemma says that the symmetric part of T_i is positive definite if the size of the subdomains, i.e. H , is sufficiently small. The proof is relatively simple, and therefore not included. The constant C appearing in the lemma depends on the coefficients $\beta(x)$ and $c(x)$ of the elliptic operator L .

LEMMA 3.3. *There exists a positive constant C , independent of the mesh parameters, such that*

$$a(u, T_i u) \geq (1 - CH)\|T_i u\|_a^2 - CH\|u\|_{a(\Omega'_i)}^2,$$

for any i , and $u \in V_h$.

The contribution from the first and zeroth order terms of the elliptic operator L is estimated in the next lemma. We prove that the contribution is of lower order in H .

LEMMA 3.4. *There exists a positive constant C , independent of the mesh parameters h and H , such that for any $i, j \neq 0$ for which Ω'_i and Ω'_j overlap*

$$1. \ s(T_i u, T_j u) \leq CH (\|T_i u\|_a^2 + \|T_j u\|_a^2)$$

$$2. \ s(T_i T_j u, T_i u) \leq CH (\|T_i u\|_a^2 + \|T_j u\|_a^2)$$

$$3. \ s(u, T_i T_j u) \leq CH \left(\|T_j u\|_a^2 + \|u\|_{a(\Omega'_i)}^2 \right)$$

for all $u \in V_h$. The same estimates hold if the bilinear form $s(\cdot, \cdot)$ is replaced by the bilinear form $c(\cdot, \cdot)$.

We leave the proof of this lemma to the interested reader. The basic idea of the proof is to use that $\|T_l u\|_{L^2(\Omega'_l)} \leq CH \|T_l u\|_{a(\Omega'_l)}$, for any $l \neq 0$. As in the previous lemma, the constants C depend on the coefficients $\beta(x)$ and $c(x)$ of the elliptic operator L . Using Lemmas 3.3 and 3.4, we now proceed to give a lower bound of the two-subdomain multiplicative Schwarz operator P_{ij} .

LEMMA 3.5. *There exists a positive constant C , independent of the mesh parameters h and H , such that for any i, j for which Ω'_i and Ω'_j overlap*

$$a(P_{ij} u, u) \geq \left(\frac{1}{2} - CH \right) \|T_i u\|_a^2 + \left(\frac{1}{2} - CH \right) \|T_j u\|_a^2 - CH \|u\|_{a(\Omega'_i)}^2,$$

for any $u \in V_h$.

Proof. We first note, by using the definition of the operators T_i and T_j and the fact that $b(\cdot, \cdot) = a(\cdot, \cdot) + s(\cdot, \cdot) + c(\cdot, \cdot)$, that

$$\begin{aligned} a(P_{ij} u, u) &= a(T_i u, u) + a(T_j u, u) - a(T_i T_j u, u) \\ &= a(T_i u, u) + a(T_j u, u) - a(T_i u, T_j u) + \\ &\quad s(T_i u, T_j u) - c(T_i u, T_j u) + 2s(T_i T_j u, T_i u) + \\ &\quad s(u, T_i T_j u) + c(u, T_i T_j u). \end{aligned}$$

The desired proof follows by using Lemmas 3.3 and 3.4. \square

We are now ready to prove the main theorem of this subsection. The upper bound is easy. It can be seen that

$$P_{ij} = T_i + (I - T_i)T_j.$$

By using the fact that $(I - T_i)$ is uniformly bounded, we obtain

$$\|P_{ij} u\|_a \leq C(\|T_i u\|_a + \|T_j u\|_a).$$

The upper bound of P can then be obtained by summing the above estimate for all possible pairs of subdomains and using Lemma 3.2. To establish the lower bound, we sum the estimate in Lemma 3.5 and use the lower bound part of Lemma 3.2, and the assumption that H is sufficiently small.

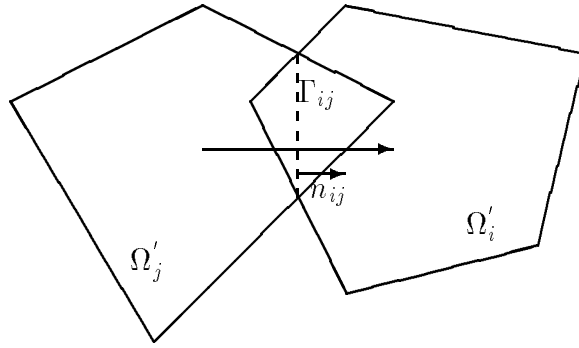


FIG. 1. The term $T_i T_j$ is kept in the Schwarz polynomial only if the flow goes from Ω'_j to Ω'_i .

3.3. A weighted local multiplicative algorithm. In this subsection, we introduce a variant of the local multiplicative algorithm that is particularly useful for fluid flow problems. The basic philosophy is the same as in the design of any upwinding type discretization schemes. We first note that the operator P has the following, more explicit, form

$$(8) \quad P = \sum_{0 \leq i \leq N} T_i - \sum_{1 \leq i \neq j \leq N} T_i T_j.$$

In other words, P is equal to the regular two-level additive Schwarz operator plus some second order perturbation terms. Since the additional second order terms enhance the nearest neighbor communication, we therefore believe they will make the overall convergence faster for the classical additive Schwarz algorithms. This observation will be confirmed by a number of numerical experiments in the next section. Borrowing a term from the Streamline Upwind Petrov-Galerkin (SUPG) methods [18, 20], the second order terms $T_i T_j$, if used properly, “stabilize” the preconditioner when solving convection-diffusion equations. The SUPG method also suggests the following version of the algorithm with weights in the upwinding directions. Let

$$(9) \quad T = \sum_{0 \leq i \leq N} T_i - \sum_{1 \leq i \neq j \leq N} \mu_{ij} T_i T_j.$$

Here μ_{ij} equals zero or μ , where $0 < \mu < 1.0$ is a constant. The choice of μ_{ij} depends on the direction of the flow. The intuition is that if the flow goes from Ω'_j to Ω'_i and if these two subdomains are neighbors, then we set μ_{ij} to be a positive constant μ , and μ_{ji} to zero. We have not exploited the possibility of using different μ_{ij} for different pairs of subdomains. Of course, if Ω'_i and Ω'_j are not neighbors we then set $\mu_{ij} = \mu_{ji} = 0$. The motivation here is exactly the same as in using the upwinding techniques in the solution of problems that involve hyperbolic components. A difference is that the usual upwinding techniques are used only at the discretization level, and our “upwinding” is introduced as a way to define the preconditioning polynomial. It is understandable that, for problems that have a strong characteristic direction, such as convection-diffusion problems, some kind of upwinding can speed up the convergence.

We now propose a heuristic method to be used to determine the flow direction. Let $\beta(x) = (b_1(x), \dots, b_d(x))^T$ be the characteristic vector of the flow. For each pair of neighboring subdomains, we choose a curve, such as Γ_{ij} in Fig. 1 or another curve in $\bar{\Omega}_i \cap \bar{\Omega}_j$,

that more or less separates the subdomains. Since we are defining preconditioners, it is not necessary to find the precise separating curve. Let n_{ij} , defined on Γ_{ij} , be the unit vector pointing from subdomain Ω'_j to Ω'_i . We define the parameters μ_{ij} by looking at the sign of a line integral

$$\mu_{ij} = \begin{cases} \mu & \text{if } \int \beta(x) \cdot n_{ij} ds > 0 \\ 0 & \text{otherwise,} \end{cases}$$

where the integral is taken along the curve Γ_{ij} .

4. Numerical experiments. In this section, we present some experimental results to numerically understand the local multiplicative Schwarz algorithms, and to compare them with the classical additive and multiplicative Schwarz algorithms for both symmetric positive definite and nonsymmetric problems. Although the proposed methods belong to the class of optimal preconditioners, some effort is needed to obtain the best performance for a particular test problem, especially in the selection of the parameter μ in both symmetric and nonsymmetric cases. We note that $\mu = 1.0$ is usually not a good choice. As mentioned earlier, our optimal convergence theory requires that the coarse grid is sufficiently fine, however, in practice, especially in the nonsymmetric cases, it is quite difficult to find a coarse grid of proper size such that the convergence is not slower than the purely local (i.e., without a coarse space) Schwarz algorithms.

We consider the following model problem on the unit square

$$\begin{cases} Lu = f & \text{in } \Omega = (0, 1) \times (0, 1), \\ u = 0 & \text{on } \partial\Omega. \end{cases}$$

The right-hand side f is always chosen such that the exact solution is $u = xe^{xy}\sin(\pi x)\sin(\pi y)$. The coefficients of L will be specified later for each test problem. We use an 256×256 uniform fine mesh throughout this section. The number of subdomains is 64 in all test cases, i.e., we use an 8×8 uniform partitioning of the domain into subdomains, with a uniform $2h$ overlap between each neighboring subdomains, where $h = 1/256$. In our experiments, the coarse grid linear system and all the subdomain linear systems are solved exactly by using a sparse linear system solver from the Argonne National Laboratory software package PETSc of Gropp and Smith [19]. All the Schwarz methods are used as left preconditioners for the CG method, or the non-restarted GMRES method, with a zero initial guess. We stop the CG or GMRES iteration as soon as the preconditioned initial residual is reduced by a factor of 10^{-5} . We discretize the PDE at both the fine and the coarse levels by the usual five-point central, or upwinding, finite difference method.

Example 0. We first test the algorithms on a simple Poisson's equation. (This is not what the new algorithm is designed for.) In Fig. 2, we show that the new algorithm is slower than the multiplicative Schwarz algorithm, but with parameter $\mu = 0.3$, faster than the additive Schwarz algorithm. Without using a proper μ , the algorithm can be slow. An 8×8 coarse solve is included in all cases. The multiplicative Schwarz algorithm is symmetrized in

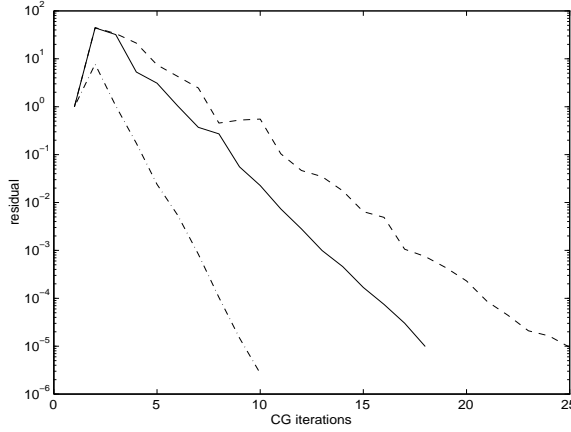


FIG. 2. The curves show the iteration history of the additive, multiplicative and the local multiplicative Schwarz preconditioned CG methods. The solid curve represents the local multiplicative Schwarz method, the dashed curve represents the additive Schwarz and the broken curve represents the multiplicative Schwarz method.

order to be able to use CG. We remark again that even though the symmetrized multiplicative Schwarz is the fastest among the three algorithms, it has the lowest parallelism. The per-step arithmetic cost of the new algorithm is higher due the repetition of the subdomain solves.

Example 1. We let $Lu = -\nabla \cdot (\nabla u) + \nabla \cdot (\beta u)$, where $\beta = (b_1, b_2)$ is a constant vector with $b_1, b_2 = 100.0$, or -100.0 . We discretize the PDE with the usual five-point *central finite difference* method. We first compare the new method, with $\mu = 0.5$, with the additive and multiplicative Schwarz methods without coarse space in the case $\beta = (100, 100)$. For the multiplicative Schwarz, we order the subdomains by the natural ordering. No coloring is incorporated in the implementation. The results are presented in the left figure of Fig. 3. It can be seen clearly that, for $\beta = (100, 100)$, the multiplicative Schwarz method is the fastest of the three. However, the situation changes, if we let $\beta = (-100, -100)$ and do not change the subdomain ordering in the multiplicative Schwarz method. As shown in the left figure of Fig. 4, the new method becomes the fastest of the three. Apparently, the changing of the flow characteristics hurts the convergence of the multiplicative Schwarz algorithm, but the new method does not suffer.

We next present cases when coarse spaces are included in the preconditioners. The optimal convergence theory for all three Schwarz algorithms requires that the coarse grid is sufficiently fine. Our numerical experiments suggest that they in fact need coarse grid of different sizes, i.e., a sufficiently fine coarse grid for one Schwarz method may not be sufficiently fine for the others. We say a coarse space is “good” if the total number of iterations is smaller than without it. A coarse grid, not sufficiently fine, usually leads to a slower convergence in all Schwarz type methods. In the right figure of Fig. 3, we present three Schwarz algorithms with three different coarse grid sizes, namely the multiplicative Schwarz with an 16×16 coarse grid; the additive Schwarz with an 32×32 coarse grid; the new method with an 64×64 coarse grid, and $\mu = 0.5$. Comparing the right figures in Fig. 3 and Fig. 4, we observe that the multiplicative Schwarz method with a coarse space of proper

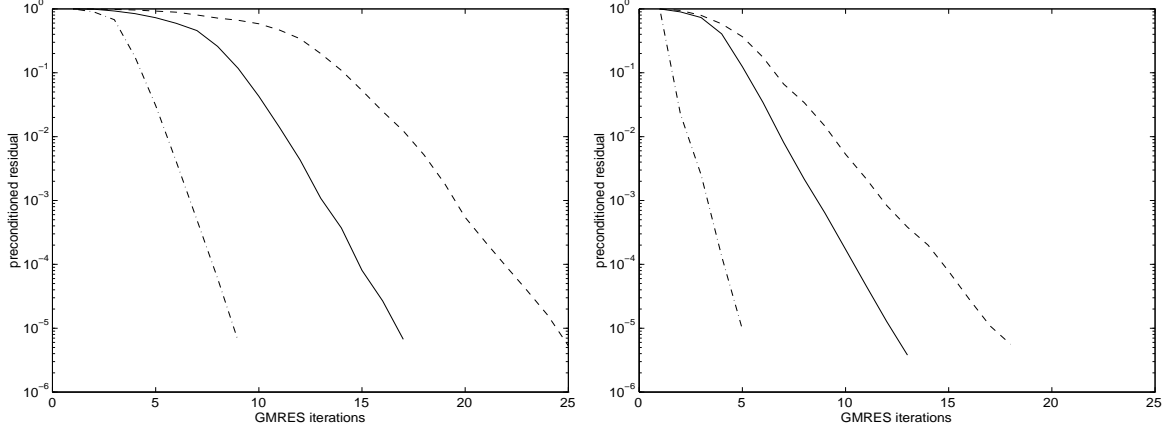


FIG. 3. Central finite difference discretization with $\beta = (100, 100)$. The left figure shows cases without coarse spaces, and the right figure shows cases with coarse spaces. The curves show the iteration history of the additive (dashed), multiplicative (broken) and the local multiplicative (solid) Schwarz preconditioned GMRES methods.

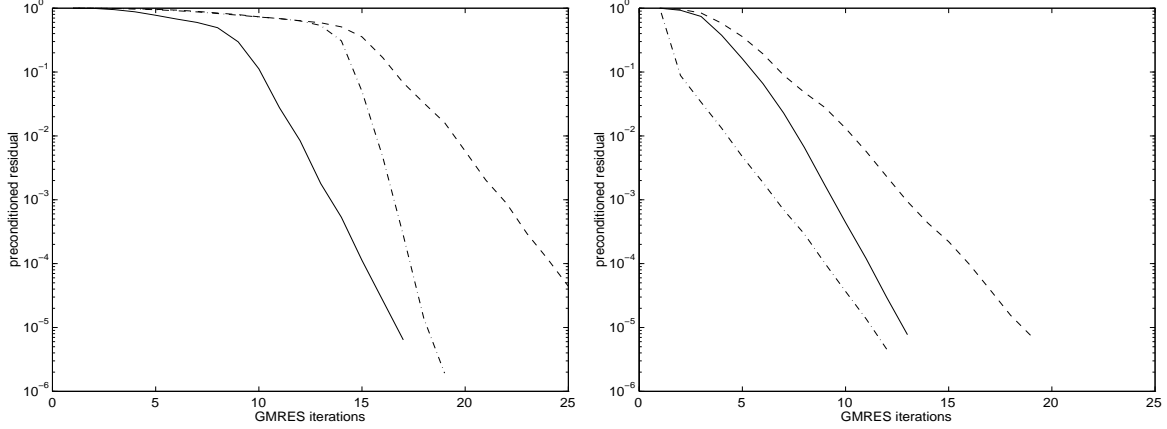


FIG. 4. Central finite difference discretization with $\beta = (-100, -100)$. The left figure shows cases without coarse spaces, and the right figure shows cases with coarse spaces. The curves show the iteration history of the additive (dashed), multiplicative (broken) and the local multiplicative (solid) Schwarz preconditioned GMRES methods.

size is always the best of the three.

Example 2. We let $Lu = -\nabla \cdot (\nabla u) + \nabla \cdot (\beta u)$, where $\beta = (b_1, b_2)$ is a constant vector with $b_1, b_2 = 1000.0$ or -1000.0 . The equation is discretized by the usual five-point *upwinding finite difference* method. We run the test code without using coarse spaces for four different constant flow directions. As before, for the multiplicative Schwarz preconditioner, we order the subdomains in the natural ordering. No coloring is assumed. For the new algorithm we use $\mu = 0.7$. The residual history is presented in Fig. 5. It is clear that if the subdomain ordering does not follow the flow characteristic direction the convergence of multiplicative Schwarz becomes significantly worse than in a case when the ordering follows the flow. Additive Schwarz is not sensitive at all to such an ordering, but is quite slow. The new algorithm does not need any special attention to the ordering, and converges faster than (a) the additive Schwarz algorithm in all four cases; (b) the worst case of the multiplicative

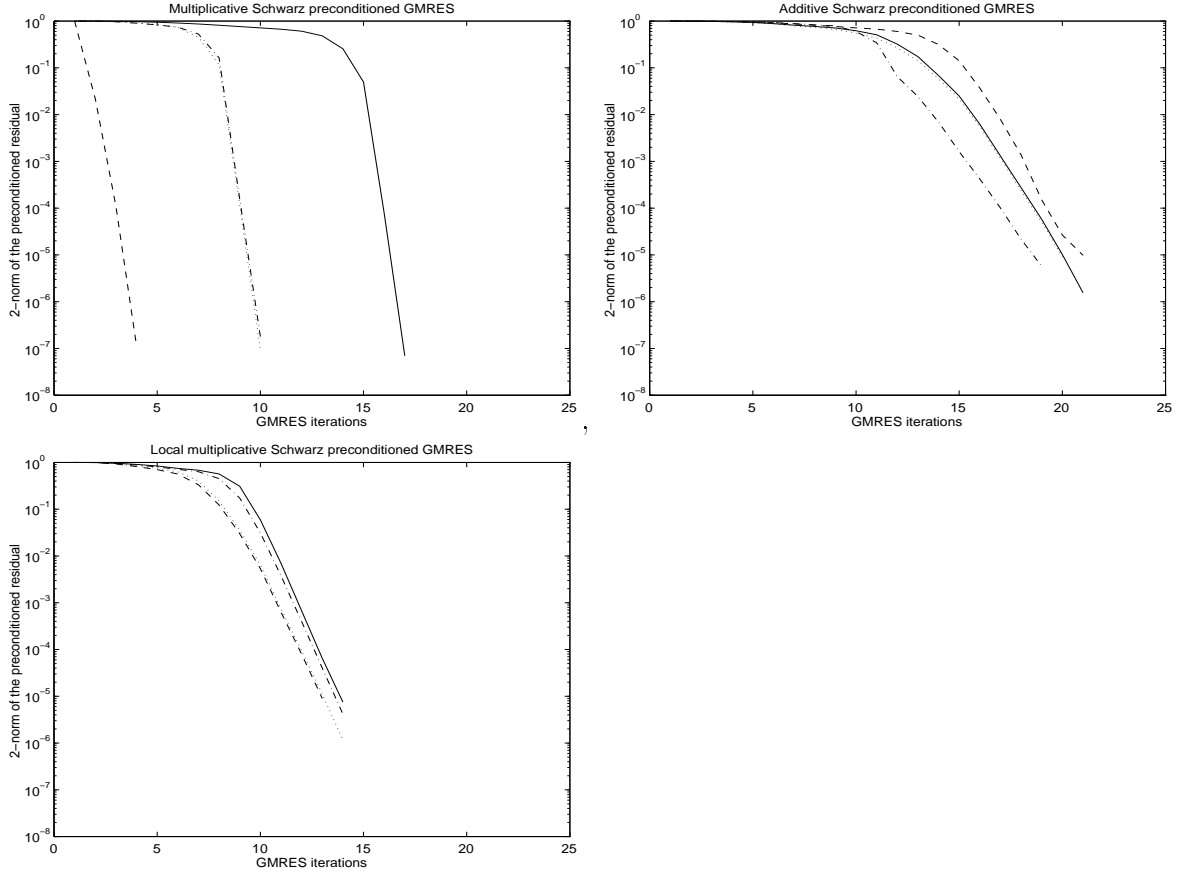


FIG. 5. The figures show the iteration history of the additive (upper left), multiplicative (upper right) and the new (lower left) Schwarz preconditioned GMRES method. The line types correspond to the flow directions, i.e., solid lines $\beta = (-1000, -1000)$, dashed lines $\beta = (1000, 1000)$, dotted lines $\beta = (1000, -1000)$ and broken lines $\beta = (-1000, 1000)$.

Schwarz algorithm.

Our experience suggests that it is by no means easy to find a coarse space of proper size in the case that the PDE is discretized by upwinding finite difference methods. Further theoretical and numerical investigation of this situation is underway.

5. Conclusion. In this paper, we introduce a new class of overlapping domain decomposition methods for solving scalar convection-diffusion problems. The method improves the classical multiplicative Schwarz methods by reducing their sensitivity with respect to the flow direction. For the Galerkin finite element discretization, we prove that the method is optimal in the sense that the convergence rate is independent of the mesh size and also the number of subdomains in both R^2 and R^3 . Numerical experiments are also reported to illustrate the rankings of the methods and some open questions are identified.

Acknowledgement. The authors are indebted to D. Keyes and O. Widlund for reading a draft of the paper and also many helpful discussions.

REFERENCES

- [1] L. M. ADAMS AND H. J. JORDAN, *Is SOR color blind?* SIAM J. Sci. Stat. Comput., 7 (1986), pp. 490–506.
- [2] X.-C. CAI, *The use of pointwise interpolation in domain decomposition methods with non-nested meshes*, SIAM J. Sci. Comput., 16 (1995), pp. 250–256.
- [3] X.-C. CAI, C. FARHAT, AND M. SARKIS, *Schwarz methods for the unsteady compressible Navier-Stokes equations on unstructured meshes*, Domain Decomposition Methods in Sciences and Engineering, R. Glowinski, J. Periaux, Z. Shi and O. Widlund, ed., 1996, Wiley (to appear).
- [4] X.-C. CAI, W. D. GROPP, D. E. KEYES, AND M. D. TIDRIRI, *Newton-Krylov-Schwarz methods in CFD*, in Proceedings of the International Workshop on the Navier-Stokes Equations, Notes in Numerical Fluid Mechanics, R. Rannacher, ed., Vieweg Verlag, Braunschweig, (1995), pp. 17–30.
- [5] X.-C. CAI AND O. B. WIDLUND, *Domain decomposition algorithms for indefinite elliptic problems*, SIAM J. Sci. Statist. Comput., 13 (1992), pp. 243–258.
- [6] ———, *Multiplicative Schwarz algorithms for nonsymmetric and indefinite elliptic problems*, SIAM J. Numer. Anal., 30 (1993), pp. 936–952.
- [7] T. F. CHAN AND T. P. MATHEW, *Domain decomposition algorithms*, Acta Num. (1994), pp. 61–143.
- [8] P. CIARLET, *The Finite Element Method for Elliptic Problems*, North-Holland, Amsterdam, 1978.
- [9] E. F. D’AZEVEDO, P. A. FORSYTH, AND W.-P. TANG, *Ordering methods for preconditioned conjugate gradient methods applied to unstructured grid problems*, SIAM J. Matrix Anal. Appl., 13 (1992), pp. 944–961.
- [10] C. FARHAT AND M. LESOINNE, *Mesh partitioning algorithms for the parallel solution of partial differential equations*, Appl. Numer. Math., 12 (1993), pp. 443–457.
- [11] M. DRYJA, M. SARKIS, AND O. B. WIDLUND, *Multilevel Schwarz methods for elliptic problems with discontinuous coefficients in three dimensions*, TR-662, Dept. of Comp. Sci., Courant Inst., 1994. (To appear in Numer. Math.)
- [12] M. DRYJA, B. F. SMITH, AND O. B. WIDLUND, *Schwarz analysis of iterative substructuring algorithms for problems in three dimensions*, SIAM J. Numer. Anal., 31 (1994), pp. 1662–1694.
- [13] M. DRYJA AND O. B. WIDLUND, *Towards a unified theory of domain decomposition algorithms for elliptic problems*, in Third International Symposium on Domain Decomposition Methods for Partial Differential Equations, T. Chan, R. Glowinski, J. P eriaux, and O. Widlund, eds., SIAM, 1990.
- [14] ———, *Domain decomposition algorithms with small overlap*, SIAM J. Sci. Comput., 15 (1994), pp. 604–620.
- [15] I. S. DUFF AND G. A. MEURANT, *The effect of ordering on preconditioned conjugate gradients*, BIT, (1989), pp. 635–657.
- [16] S. C. EISENSTAT, H. C. ELMAN, AND M. H. SCHULTZ, *Variational iterative methods for nonsymmetric systems of linear equations*, SIAM J. Numer. Anal., 20 (1983), pp. 345–357.
- [17] H. C. ELMAN AND M. P. CHERNESKY, *Ordering effects on relaxation methods applied to the discrete one-dimensional convection-diffusion equation*, SIAM J. Numer. Anal., 30 (1993), pp. 1268–1290.
- [18] L. P. FRANCA, S. R. FREY, AND T. J. R. HUGHES, *Stabilized finite element methods: I. Application to the advection-diffusion model*, Comp. Meths. Appl. Mech. Eng., 95 (1992), pp. 253–276.
- [19] W. D. GROPP AND B. F. SMITH, *Scalable, extensible, and portable numerical libraries*, Proceedings of Scalable Parallel Libraries Conference, IEEE, 1993, pp. 87–93.
- [20] C. JOHNSON, *Numerical Solution of Partial Differential Equations by the Finite Element Method*, Cambridge University Press, 1987.
- [21] Y. SAAD AND M. H. SCHULTZ, *GMRES: A generalized minimum residual algorithm for solving non-symmetric linear systems*, SIAM J. Sci. Stat. Comput., 7 (1986), pp. 856–869.
- [22] H. A. SCHWARZ, *Gesammelte Mathematische Abhandlungen*, vol. 2, Springer, Berlin, 1890, pp. 133–143. First published in Vierteljahrsschrift der Naturforschenden Gesellschaft in Z urich, volume 15, 1870, pp. 272–286.
- [23] B. F. SMITH, P. E. BJ RSTAD, AND W. D. GROPP, *Domain Decomposition: Parallel Multilevel Methods for Elliptic Partial Differential Equations*, Cambridge University Press, 1995. To appear.
- [24] V. VENKATAKRISHNAN, *Implicit schemes and parallel computing in unstructured grid CFD*, ICASE Report No. 95-28, 1995.